

# SMOOTHNESS OF DENSITY AND ERGODICITY FOR STATE-DEPENDENT SWITCHING DIFFUSIONS

YAOZHONG HU, DAVID NUALART, AND XIAOBIN SUN AND YINGCHAO XIE

**ABSTRACT.** This paper is concerned with a class of stochastic differential equations with state-dependent switching. The Malliavin calculus is used to study the smoothness of the density of the solutions under the Hörmander type conditions. Moreover, the strong Feller property of the process is obtained by using the Bismut formula. The irreducibility of the semigroup associated with the equations is discussed under some natural conditions. As a consequence the existence and uniqueness of the invariant measure and then the ergodicity for the equations are also discussed.

## 1. INTRODUCTION

This paper considers the following state-dependent switching diffusion process on  $\mathbb{R}^n$ :

$$dX_t = b(X_t, \alpha_t)dt + \sigma(X_t, \alpha_t)dW_t, \quad (X_0, \alpha_0) = (x, \alpha) \in \mathbb{R}^n \times \mathbb{S}, \quad (1.1)$$

where  $\mathbb{S} = \{1, 2, \dots, m_0\}$  and  $\{\alpha_t, t \geq 0\}$  is an  $\mathbb{S}$ -valued switching process described by

$$\mathbb{P}\{\alpha_{t+\Delta} = j | \alpha_t = i, X_s, \alpha_s, s \leq t\} = \begin{cases} q_{ij}(X_t)\Delta + o(\Delta), & i \neq j \\ 1 + q_{ii}(X_t)\Delta + o(\Delta), & i = j, \end{cases} \quad (1.2)$$

and  $Q(x) = (q_{ij}(x))_{1 \leq i, j \leq m_0}$  is a  $Q$ -matrix dependent on the state  $x$  of the process  $X_t$ .

Switching diffusions of this type have been studied in literature and have potential applications in mathematical finance. Existence and uniqueness of the solution, Feller property, ergodicity and other important properties have been studied. In [12, 14], the authors compare the processes  $(X_t, \alpha_t)$  with an auxiliary state *independent* switching diffusion  $(V_t, \psi_t)$ , which satisfy equations similar to (1.1) and (1.2) with  $X_t$  replaced by  $V_t$  and  $\alpha_t$  by  $\psi_t$ , but the corresponding  $Q$  matrix satisfies  $q_{ij}(x) = 1$  when  $i \neq j$ . Combining some estimates of the Radon-Nikodym derivative of the law of  $(X_t, \alpha_t)$  with respect to the law of  $(V_t, \psi_t)$  and the estimates of the process  $(V_t, \psi_t)$ , the Feller and the strong Feller properties of  $(X_t, \alpha_t)$  are obtained under Lipschitz and uniform boundedness conditions on  $q_{ij}(x)$ . Let us

---

*Key words and phrases.* Malliavin calculus, state-dependent switching diffusions, smoothness of density, Bismut formula, strong Feller property, irreducibility, invariant measure, ergodicity.

Y. Hu is partially supported by a grant from the Simons Foundation #209206 and by a General Research Fund of University of Kansas.

D. Nualart is supported by the NSF grant DMS1208625.

X. Sun and Y. Xie are supported by the National Natural Science Foundation of China (11271169) and the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

also mention that in [15], the strong Feller property is obtained under Hölder and uniform boundedness conditions of the  $Q$ -matrix  $q_{ij}(x)$ .

In this paper we derive the strong Feller property of the process  $(X_t, \alpha_t)$  without assuming the Hölder continuity of the  $Q$ -matrix  $q_{ij}(x)$ . We only assume that  $q_{ij}(x)$  is uniformly bounded. Our approach is completely different from the methodology used previously and it is based on Malliavin calculus. More precisely, we establish a Bismut formula for state-dependent switching diffusions which allows us to derive the strong Feller property.

To establish a Bismut formula, we need to develop the Malliavin calculus for the state-dependent switching diffusions  $X_t$  satisfying (1.1) and (1.2). The difficulty here is the appearance of the Markov switching terms  $\alpha(t)$ , which is a jump processes. Our procedure is to perform perturbations of the underlying Brownian motion, keeping the Poisson random measure which drives the switching unperturbed. The technique for this analysis is inspired in the partial Malliavin calculus which can be regarded as a stochastic calculus of variation for random variables with values in a Hilbert space.

After developing the Malliavin calculus for state-dependent switching diffusions, we investigate the smoothness of the density of the law of the random vector  $X_t$ , for a fixed  $t > 0$ , with respect to the Lebesgue measure. In order to establish this smoothness of density property we need to show that the determinant of the Malliavin matrix of  $X_t$  has negative moments of all orders. In the classical diffusion case, this is guaranteed by a Hörmander-type non degeneracy condition. To follow this classical approach we immediately encounter a difficulty: the process  $X_t$  depends on the discrete process  $\alpha_t$ , and the application of Itô's formula yields some jump terms. This requires a version of the classical Norris lemma in random intervals, which works only if the length of the interval is bounded below by a constant. However, in general this is not true. To overcome this difficulty we shall use the following strategy inspired by [5]. First we notice that the jump times form a subset of the jump times of some Poisson process  $N_t$ , independent of the driving Brownian motion  $W_t$ . Then conditioning on  $N_t = k$ , there exists a random interval  $[T_1, T_2]$  such that  $T_2 - T_1 \geq \frac{t}{k+1}$ . On this random time interval, the Itô's formula for  $(X_t, \alpha_t)$  will not produce a jump term, and we can apply the classical procedure. Details is given in section 4.

The paper is organized as follows. In the next section, we introduce some notation and assumptions that we use throughout the paper. We develop the Malliavin calculus for switching diffusions in Section 3. In Section 4, we show that the determinant of the Malliavin covariance matrix has all negative finite moments under a suitable uniform Hörmander's condition. We establish the Bismut formula in Section 5 which can be used to obtain the strong Feller property in Section 7. In Section 6, the existence of invariant measure is obtained. Finally, Section 7 is devoted to study the ergodicity by means of the strong Feller property and the irreducibility of the process  $(X_t, \alpha_t)$ .

## 2. PRELIMINARIES

Fix a time interval  $[0, T]$ . Let  $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$  be the  $d$ -dimensional canonical Wiener space with the natural filtration  $\mathcal{F}_1 = \{\mathcal{F}_1(t), 0 \leq t \leq T\}$ . That is,  $\Omega_1$  is the set of all continuous maps  $\omega_1 : [0, T] \rightarrow \mathbb{R}^d$  such that  $\omega_1(0) = 0$  and  $\mathcal{F}_1$  is the completion of the Borel  $\sigma$  field of  $\Omega_1$  with respect to  $\mathbb{P}_1$ , where  $\mathbb{P}_1$  is the canonical Wiener

measure. Then,  $W = \{W_t(\omega_1) := \omega_1(t), t \in [0, T]\}$  is a  $d$ -dimensional Brownian motion.

Let  $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$  be another complete probability space with a filtration  $\mathcal{F}_2 = \{\mathcal{F}_2(t), t \in [0, T]\}$  satisfying the usual conditions, on which  $N(dt, dz)$  is a Poisson random measure on  $\mathbb{R}_+ \times \mathbb{R}$  whose intensity is the Lebesgue measure. Denote the product probability space by  $(\Omega, \mathcal{F}, \mathbb{P}) := (\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mathbb{P}_1 \times \mathbb{P}_2)$  with the product filtration  $\mathcal{F} = \{\mathcal{F}_t, t \in [0, T]\}$ , where  $\mathcal{F}_t = \mathcal{F}_1(t) \times \mathcal{F}_2(t)$ . We extend  $W_t$  and  $N(A)$ , where  $A$  is Borel set in  $\mathbb{R}_+ \times \mathbb{R}$  with finite Lebesgue measure, to random variables defined on  $\Omega$  by letting  $W_t(\omega) = \omega_1(t)$  and  $N(A, \omega) = N(A, \omega_2)$ , respectively, if  $\omega = (\omega_1, \omega_2)$ . Notice that on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  $W$  and  $N$  are independent.

Let  $\mathbb{S} = \{1, 2, \dots, m_0\}$ , where  $m_0$  is a given positive integer which will be fixed throughout the paper. Let  $Q(x) = (q_{ij}(x))_{1 \leq i, j \leq m_0}$  be a  $Q$ -matrix for any  $x \in \mathbb{R}^n$ , satisfying the following assumption:

- (i)  $q_{ij}(x)$  is Borel measurable and exists  $K > 0$  such that  $\sup_{i, j \in \mathbb{S}, x \in \mathbb{R}^n} |q_{ij}(x)| \leq K$ ,
- (ii)  $q_{ij}(x) \geq 0$  for  $x \in \mathbb{R}^n$  and  $i \neq j$ ,
- (iii)  $q_{ii}(x) = -\sum_{j \neq i} q_{ij}(x)$  for  $x \in \mathbb{R}^n$  and  $i \in \mathbb{S}$ .

Suppose that  $b : \mathbb{R}^n \times \mathbb{S} \rightarrow \mathbb{R}^n$  and  $\sigma : \mathbb{R}^n \times \mathbb{S} \rightarrow \mathbb{R}^{nd}$  be given mappings satisfying some properties which will be specified later. We consider the following state-dependent switching diffusion process on  $\mathbb{R}^n$  with switching rate  $Q$

$$dX_t = b(X_t, \alpha_t)dt + \sigma(X_t, \alpha_t)dW_t, \quad (X_0, \alpha_0) = (x, \alpha) \in \mathbb{R}^n \times \mathbb{S}, \quad (2.1)$$

where  $\{\alpha_t, t \geq 0\}$  is an  $\mathbb{S}$ -valued Markov process described by

$$\mathbb{P}\{\alpha_{t+\Delta} = j | \alpha_t = i, X_s, \alpha_s, s \leq t\} = \begin{cases} q_{ij}(X_t)\Delta + o(\Delta), & i \neq j \\ 1 + q_{ii}(X_t)\Delta + o(\Delta), & i = j. \end{cases} \quad (2.2)$$

It is well known (see [13, 15]) that the process  $\{\alpha_t, 0 \leq t \leq T\}$  can be described in the following manner by using the Poisson random measure  $N$ . Introduce the function  $g : \mathbb{R}^n \times \mathbb{S} \times [0, m_0(m_0 - 1)K] \rightarrow \mathbb{R}$  defined by

$$g(x, i, z) = \sum_{j \in \mathbb{S} \setminus i} (j - i)1_{z \in \Delta_{ij}(x)}, \quad \forall i \in \mathbb{S},$$

where  $\Delta_{ij}(x)$  are the consecutive (with respect to the lexicographic ordering on  $\mathbb{S} \times \mathbb{S}$ ) left-closed, right-open intervals of  $\mathbb{R}_+$ , each having length  $q_{ij}(x)$ , with  $\Delta_{12}(x) = [0, q_{12}(x))$ . Then, Equation (2.2) can also be written as

$$d\alpha(t) = \int_{[0, m_0(m_0 - 1)K]} g(X_{t-}, \alpha_{t-}, z)N(dt, dz). \quad (2.3)$$

For  $k \in \mathbb{N}$  we denote by  $C^k(\mathbb{R}^n \times \mathbb{S}; \mathbb{R}^n)$  the family of all  $\mathbb{R}^n$ -valued functions  $f(x, \alpha)$  on  $\mathbb{R}^n \times \mathbb{S}$  which are  $k$ -times continuously differentiable in  $x$  for any  $\alpha \in \mathbb{S}$ . The  $k$ -th derivative tensor of  $f$  with respect to  $x$  is denoted by  $\nabla^k f(x, \alpha)$ .

We denote by  $|\cdot|$  the Euclidean norm. We consider the metric  $\Lambda$  on  $\mathbb{R}^n \times \mathbb{S}$  given by  $\Lambda((x, i), (y, j)) = |x - y| + d(i, j)$ , for  $x, y \in \mathbb{R}^n, i, j \in \mathbb{S}$ , where  $d(i, j) = 0$  if  $i = j$  and  $d(i, j) = 1$  if  $i \neq j$ . Let  $\mathcal{B}_b(\mathbb{R}^n \times \mathbb{S})$  be the family of all bounded Borel measurable functions on  $\mathbb{R}^n \times \mathbb{S}$ .

We will make use of the following assumptions on the functions  $b$  and  $\sigma$ :

**(H1)** There is a positive constant  $C_1$  such that

$$|b(x, i) - b(y, i)| \vee |\sigma(x, i) - \sigma(y, i)| \leq C_1 |x - y| \quad \text{for any } x, y \in \mathbb{R}^n, i \in \mathbb{S}.$$

**(H<sub>k</sub>)** Fix an integer  $k \geq 1$ . For each  $i = 1, \dots, d$ , the functions  $b$  and  $\sigma_i$  belong to  $C^k(\mathbb{R}^n \times \mathbb{S}; \mathbb{R}^n)$ , and have bounded partial derivatives up to the order  $k$ .

**(H<sub>∞</sub>)** For each  $i = 1, \dots, d$  and for any  $k \geq 1$ , the functions  $b$  and  $\sigma_i$  belong to  $C^k(\mathbb{R}^n \times \mathbb{S}; \mathbb{R}^n)$ , and have bounded partial derivatives of all orders.

It is clear that **(H<sub>k</sub>)** implies **(H1)** for any  $k \geq 1$ .

Under condition **(H1)**, the system of equations (2.1) and (2.3) possesses a unique strong solution  $\{(X_t, \alpha_t), t \in [0, T]\}$ . Moreover, for any  $p \geq 2$ ,  $\mathbb{E}(\sup_{0 \leq t \leq T} |X_t(x, \alpha)|^p) \leq C_2$ , where  $C_2$  is a positive constant depending on  $p, T$  and  $x$ . The solution is a Markov process and the associated Markov semigroup  $P_t$  satisfies

$$P_t f(x, \alpha) = \mathbb{E} f(X_t(x, \alpha), \alpha_t(x, \alpha)), \quad t \in [0, T], f \in \mathcal{B}_b(\mathbb{R}^n \times \mathbb{S}).$$

We refer, for instance, to [16] for a detailed presentation and proofs of the above results.

Along the paper  $C$  will denote a generic constant which may vary from line to line and it might depend on  $T$ , the exponent  $p \geq 2$ , the initial condition  $x$  and a fixed element  $h \in H$ .

### 3. THE MALLIAVIN CALCULUS FOR STATE DEPENDENT SWITCHING DIFFUSION

In this section we analyze the regularity, in the sense of Malliavin calculus, of the solution  $X_t$  to the system (2.1) and (2.3). The procedure is to perform perturbations of the underlying Brownian motion, keeping the Poisson random measure invariant. The technique for this analysis is inspired in the partial Malliavin calculus which can be regarded as a stochastic calculus of variation for random variables with values on the Hilbert space  $L^2(\Omega_2)$ .

We follow the notation introduced in the Preliminaries. Denote by  $H$  the Hilbert space  $H = L^2([0, T]; \mathbb{R}^d)$ , equipped with the inner product  $\langle h_1, h_2 \rangle_H = \int_0^T \langle h_1(s), h_2(s) \rangle_{\mathbb{R}^d} ds$ .

For a Hilbert space  $U$  and a real number  $p \geq 1$ , we denote by  $L^p(\Omega_1; U)$  the space of  $U$ -valued random variables  $\xi$  such that  $\mathbb{E}_1 \|\xi\|_U^p < \infty$ , where  $\mathbb{E}_1$  is the mathematical expectation on the probability space  $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ . We also set  $L^{\infty-}(\Omega_1; U) := \cap_{p < \infty} L^p(\Omega_1; U)$ .

We introduce the derivative operator for a random variable  $F$  in the space  $L^{\infty-}(\Omega_1; U)$  following the approach of Malliavin in [7]. We say that  $F$  belongs to  $\mathbb{D}^{1,\infty}(U)$  if there exists  $DF \in L^{\infty-}(\Omega_1; H \otimes U)$  such that for any  $h \in H$ ,

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}_1 \left\| \frac{F(\omega_1 + \varepsilon \int_0^\cdot h_s ds) - F(\omega_1)}{\varepsilon} - \langle DF, h \rangle_H \right\|_U^p = 0$$

holds for every  $p \geq 1$ . In this case, we define the Malliavin derivative of  $F$  in the direction  $h$  by  $D^h F := \langle DF, h \rangle_H$ . Then, for any  $p \geq 1$  we define the Sobolev space  $\mathbb{D}^{1,p}(U)$  as the completion of  $\mathbb{D}^{1,\infty}(U)$  under the following norm

$$\|F\|_{1,p,U} = [\mathbb{E}_1 \|F\|_U^p]^{1/p} + [\mathbb{E}_1 \|DF\|_{H \otimes U}^p]^{1/p}.$$

By induction we define the  $k$ th derivative by  $D^k F = D(D^{k-1} F)$ , which is a random element with values in  $H^{\otimes k} \otimes U$ . For any integer  $k \geq 1$ , the Sobolev space  $\mathbb{D}^{k,p}(U)$  is the completion of  $\mathbb{D}^{k,\infty}(U)$  under the norm

$$\|F\|_{k,p,U} = \|F\|_{k-1,p,U} + \|D^k F\|_{1,p,H^{\otimes k} \otimes U}.$$

It turns out that  $D$  is a closed operator from  $L^p(\Omega_1; U)$  to  $L^p(\Omega_1; H \otimes U)$ . Its adjoint  $\delta$  is called the divergence operator, and is continuous from  $\mathbb{D}^{1,p}(H \otimes U)$  to  $L^p(\Omega_1; U)$  for any  $p > 1$ . The duality relationship reads

$$\mathbb{E}_1(\langle DF, u \rangle_{H \otimes U}) = \mathbb{E}_1(\langle F, \delta(u) \rangle_U),$$

for any  $F \in \mathbb{D}^{1,p}(U)$  and  $u \in \mathbb{D}^{1,q}(H \otimes U)$ , with  $\frac{1}{p} + \frac{1}{q} = 1$ .

A square integrable random variable  $F \in L^2(\Omega)$  can be identified with an element of  $L^2(\Omega_1; V)$ , where  $V = L^2(\Omega_2)$ .

The following is the main result of this section.

**Theorem 3.1.** *Suppose that Hypothesis  $(\mathbf{H}_2)$  holds. Then for any  $t \in [0, T]$  and any  $h \in H$ ,  $X_t \in \mathbb{D}^{1,\infty}(\mathbb{R}^n \otimes V)$  and  $D^h X_t$  satisfies*

$$\begin{cases} dD^h X_t = \nabla b(X_t, \alpha_t) D^h X_t dt + \sum_{i=1}^d \nabla \sigma_i(X_t, \alpha_t) D^h X_t dW_t^i + \sigma(X_t, \alpha_t) h_t dt, \\ D^h X_0 = 0. \end{cases} \quad (3.1)$$

To prove the theorem, let  $\{(X_t^{\varepsilon h}(x, \alpha), \alpha_t^{\varepsilon h}(x, \alpha)), t \in [0, T]\}$  be the solution of equations (2.1) and (2.3) with  $W_t$  replaced by  $W_t + \varepsilon \int_0^t h_s ds$ , where  $\varepsilon \in (0, 1)$ , that is, (2.3) would be

$$\begin{cases} dX_t^{\varepsilon h} = b(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) dt + \sigma(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) dW_t + \varepsilon \sigma(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) h_t dt, \\ (X_0^{\varepsilon h}, \alpha_0^{\varepsilon h}) = (x, \alpha) \in \mathbb{R}^n \times \mathbb{S} \end{cases} \quad (3.2)$$

and

$$\mathbb{P}\{\alpha_{t+\Delta}^{\varepsilon h} = j | \alpha_t^{\varepsilon h} = i, X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}, s \leq t\} = \begin{cases} q_{ij}(X_t^{\varepsilon h}) \Delta + o(\Delta), & i \neq j \\ 1 + q_{ii}(X_t^{\varepsilon h}) \Delta + o(\Delta), & i = j. \end{cases} \quad (3.3)$$

Then, we can write

$$\begin{aligned} \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} &= \frac{1}{\varepsilon} \int_0^t [b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_s, \alpha_s)] ds \\ &\quad + \frac{1}{\varepsilon} \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - \sigma(X_s, \alpha_s)] dW_s + \int_0^t \sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) h_s ds \\ &= \frac{1}{\varepsilon} \int_0^t [b(X_s^{\varepsilon h}, \alpha_s) - b(X_s, \alpha_s)] ds + \frac{1}{\varepsilon} \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s) - \sigma(X_s, \alpha_s)] dW_s \\ &\quad + \int_0^t \sigma(X_s^{\varepsilon h}, \alpha_s) h_s ds + \phi_t^\varepsilon, \end{aligned}$$

where

$$\begin{aligned} \phi_t^\varepsilon &= \frac{1}{\varepsilon} \int_0^t [b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_s^{\varepsilon h}, \alpha_s)] ds + \frac{1}{\varepsilon} \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - \sigma(X_s^{\varepsilon h}, \alpha_s)] dW_s \\ &\quad + \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - \sigma(X_s^{\varepsilon h}, \alpha_s)] h_s ds. \end{aligned}$$

To show Theorem 3.1 we first show the convergence  $\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |\phi_t^\varepsilon|^p \right] = 0$  for all  $p \geq 2$ . To this end we need some preliminary lemmas.

**Lemma 3.2.** *Suppose that Hypothesis (H1) holds. Then for any  $h \in H$  and  $p \geq 2$ , we have*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^{\varepsilon h}|^p \right] \leq C.$$

*Proof* From Equation (3.2) it is easy to see that

$$\begin{aligned} |X_t^{\varepsilon h}|^p &\leq C \left[ |x|^p + \left| \int_0^t b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) ds \right|^p + \left| \int_0^t \sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) dW_s \right|^p \right. \\ &\quad \left. + \varepsilon^p \left| \int_0^t \sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) h_s ds \right|^p \right] := C \left[ |x|^p + I_1(t) + I_2(t) + I_3(t) \right]. \end{aligned}$$

By Hölder's and Burkholder-Davis-Gundy's inequalities, we obtain

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} (I_1(t) + I_2(t) + I_3(t)) \right] \leq C \int_0^T (\mathbb{E} |X_s^{\varepsilon h}|^p + 1) ds.$$

Then the desired estimate follows from Gronwall's lemma. ■

**Lemma 3.3.** *Suppose that Hypothesis (H1) holds. Then for any  $h \in H$ ,  $p \geq 2$  and  $0 \leq s \leq t \leq T$ , we have*

$$\mathbb{E} |X_t^{\varepsilon h} - X_s^{\varepsilon h}|^p \leq C(t-s)^{p/2}.$$

*Proof* Note that for  $0 \leq s \leq t \leq T$ ,

$$X_t^{\varepsilon h} - X_s^{\varepsilon h} = \int_s^t b(X_r^{\varepsilon h}, \alpha_r^{\varepsilon h}) dr + \int_s^t \sigma(X_r^{\varepsilon h}, \alpha_r^{\varepsilon h}) dW_r + \varepsilon \int_s^t \sigma(X_r^{\varepsilon h}, \alpha_r^{\varepsilon h}) h_r dr.$$

Then, we have

$$\begin{aligned} |X_t^{\varepsilon h} - X_s^{\varepsilon h}|^p &\leq C \left[ \left| \int_s^t b(X_r^{\varepsilon h}, \alpha_r^{\varepsilon h}) dr \right|^p + \left| \int_s^t \sigma(X_r^{\varepsilon h}, \alpha_r^{\varepsilon h}) dW_r \right|^p \right. \\ &\quad \left. + \varepsilon^p \left| \int_s^t \sigma(X_r^{\varepsilon h}, \alpha_r^{\varepsilon h}) h_r dr \right|^p \right]. \end{aligned}$$

Burkholder-Davis-Gundy's and Hölder's inequalities yield

$$\mathbb{E} |X_t^{\varepsilon h} - X_s^{\varepsilon h}|^p \leq C \mathbb{E} \left[ \sup_{0 \leq r \leq T} (1 + |X_r^{\varepsilon h}|^p) \right] \left( (t-s)^p + (t-s)^{p/2} + \|h\|_H^p (t-s)^{p/2} \right).$$

By Lemma 3.2 we can write

$$\mathbb{E} |X_t^{\varepsilon h} - X_s^{\varepsilon h}|^p \leq C(t-s)^{p/2},$$

which yields the result. ■

The following will be a basic ingredient in the proof of Theorem 3.1.

**Lemma 3.4.** *Suppose that Hypothesis (H1) holds. Then for any  $h \in H$  and  $p \geq 2$  we have*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |\phi_t^\varepsilon|^p \right] = 0.$$

*Proof* By Hölder's and Burkholder-Davis-Gundy's inequalities, we can write

$$\begin{aligned}
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\phi_t^\varepsilon|^p \right] &\leq \frac{C}{\varepsilon^p} \mathbb{E} \int_0^T |b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_s^{\varepsilon h}, \alpha_s)|^p ds \\
&\quad + \frac{C(1 + \|h\|_H^p)}{\varepsilon^p} \mathbb{E} \int_0^T |\sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - \sigma(X_s^{\varepsilon h}, \alpha_s)|^p ds \\
&:= \frac{C}{\varepsilon^p} (I_4 + I_5). \tag{3.4}
\end{aligned}$$

We will deal with each of the above terms separately. Consider a uniform partition of the interval  $[0, T]$  into  $M$  subintervals and set  $\eta = \frac{T}{M}$ . The term  $I_4$  can be bounded as follows

$$\begin{aligned}
I_4 &\leq \mathbb{E} \sum_{k=0}^M \int_{k\eta}^{(k+1)\eta} |b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_s^{\varepsilon h}, \alpha_s)|^p ds \\
&\leq C \sum_{k=0}^M \left[ \mathbb{E} \int_{k\eta}^{(k+1)\eta} |b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_{k\eta}^{\varepsilon h}, \alpha_s^{\varepsilon h})|^p ds \right. \\
&\quad + \mathbb{E} \int_{k\eta}^{(k+1)\eta} |b(X_{k\eta}^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_{k\eta}^{\varepsilon h}, \alpha_s)|^p ds \\
&\quad \left. + \mathbb{E} \int_{k\eta}^{(k+1)\eta} |b(X_{k\eta}^{\varepsilon h}, \alpha_s) - b(X_s^{\varepsilon h}, \alpha_s)|^p ds \right] \\
&:= C \sum_{k=0}^M [I_{41k} + I_{42k} + I_{43k}]. \tag{3.5}
\end{aligned}$$

For  $k = 0, \dots, M$  applying Hypothesis **(H1)** and Lemma 3.3, we get

$$\begin{aligned}
I_{41k} &\leq C \int_{k\eta}^{(k+1)\eta} \mathbb{E} |X_s^{\varepsilon h} - X_{k\eta}^{\varepsilon h}|^p ds \\
&\leq C \int_{k\eta}^{(k+1)\eta} (s - k\eta)^{p/2} ds \leq C\eta^{p/2+1}. \tag{3.6}
\end{aligned}$$

Likewise, for  $k = 0, \dots, M$ , we have

$$I_{43k} \leq C\eta^{p/2+1}. \tag{3.7}$$

The term  $I_{42k}$  requires a more careful analysis. First we make the following decomposition

$$\begin{aligned}
I_{42k} &\leq C \mathbb{E} \int_{k\eta}^{(k+1)\eta} |b(X_{k\eta}^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_{k\eta}^{\varepsilon h}, \alpha_{k\eta}^{\varepsilon h})|^p ds \\
&\quad + C \mathbb{E} \int_{k\eta}^{(k+1)\eta} |b(X_{k\eta}^{\varepsilon h}, \alpha_{k\eta}^{\varepsilon h}) - b(X_{k\eta}^{\varepsilon h}, \alpha_s)|^p ds := C (I_{421k} + I_{422k}). \tag{3.8}
\end{aligned}$$

Applying Lemma 3.2 and the boundedness of  $Q(x)$ , the term  $I_{421k}$  in (3.8) can be estimated as follows

$$\begin{aligned}
I_{421k} &= \mathbb{E} \sum_{i \in \mathbb{S}} \sum_{j \neq i} \int_{k\eta}^{(k+1)\eta} |b(X_{k\eta}^{\varepsilon h}, j) - b(X_{k\eta}^{\varepsilon h}, i)|^p 1_{\{\alpha_s^{\varepsilon h} = j\}} 1_{\{\alpha_{k\eta}^{\varepsilon h} = i\}} ds \\
&\leq C \mathbb{E} \sum_{i \in \mathbb{S}} \sum_{j \neq i} \int_{k\eta}^{(k+1)\eta} [1 + |X_{k\eta}^{\varepsilon h}|^p] 1_{\{\alpha_{k\eta}^{\varepsilon h} = i\}} \mathbb{E}[1_{\{\alpha_s^{\varepsilon h} = j\}} | X_{k\eta}^{\varepsilon h}, \alpha_{k\eta}^{\varepsilon h} = i] ds \\
&\leq C \mathbb{E} \sum_{i \in \mathbb{S}} \int_{k\eta}^{(k+1)\eta} [1 + |X_{k\eta}^{\varepsilon h}|^p] 1_{\{\alpha_{k\eta}^{\varepsilon h} = i\}} \left[ \sum_{j \neq i} q_{ij}(X_{k\eta}^{\varepsilon h})(s - k\eta) + o(s - k\eta) \right] ds \\
&\leq C \int_{k\eta}^{(k+1)\eta} O(\eta) ds \leq C\eta^2. \tag{3.9}
\end{aligned}$$

To handle the term  $I_{422k}$  in (3.8) we will use a coupling technique for the Markov switching process  $\alpha_t$  (see [3, p. 11]) which we recall here. Let  $\{(\alpha_t, \alpha_t^{\varepsilon h}), t \geq 0\}$  be a discrete random process with a finite-state space  $\mathbb{S} \times \mathbb{S}$  such that

$$\begin{aligned}
&\mathbb{P}\{(\alpha_{t+\Delta}, \alpha_{t+\Delta}^{\varepsilon h}) = (i, j) | (\alpha_t, \alpha_t^{\varepsilon h}) = (k, l), X_t, X_t^{\varepsilon h}\} \\
&= \begin{cases} q_{(k,l)(i,j)}(X_t, X_t^{\varepsilon h})\Delta + o(\Delta), & (k, l) \neq (i, j) \\ 1 + q_{(k,l)(k,l)}(X_t, X_t^{\varepsilon h})\Delta + o(\Delta), & (k, l) = (i, j). \end{cases}
\end{aligned}$$

In the above expression the matrix  $(q_{(k,l)(i,j)}(x, x^{\varepsilon h}))$  is the basic coupling of  $Q(x) = (q_{kl}(x))$  and  $Q(x^{\varepsilon h}) = (q_{ij}(x^{\varepsilon h}))$  satisfying

$$\begin{aligned}
Q(x, x^{\varepsilon h})f(k, l) &:= \sum_{(i,j) \in \mathbb{S} \times \mathbb{S}} q_{(k,l)(i,j)}(x, x^{\varepsilon h})(f(i, j) - f(k, l)) \\
&= \sum_i (q_{ki}(x) - q_{li}(x^{\varepsilon h}))^+ (f(i, l) - f(k, l)) \\
&\quad + \sum_i (q_{li}(x^{\varepsilon h}) - q_{ki}(x))^+ (f(k, i) - f(k, l)) \\
&\quad + \sum_i (q_{ki}(x) \wedge q_{li}(x^{\varepsilon h}))(f(i, i) - f(k, l)),
\end{aligned}$$

where  $f(\cdot, \cdot)$  is defined on  $\mathbb{S} \times \mathbb{S}$  and where we denote  $a \wedge b = \min\{a, b\}$  and  $a^+ = \max\{a, 0\}$ . With this coupling, for  $i_1, i, j, l \in \mathbb{S}$  with  $i \neq j$  and  $s \in [k\eta, (k+1)\eta)$ , we have

$$\begin{aligned}
&\mathbb{E}[1_{\{\alpha_s = j\}} | \alpha_{k\eta} = i_1, \alpha_{k\eta}^{\varepsilon h} = i, X_{k\eta}, X_{k\eta}^{\varepsilon h}] \\
&= \sum_{l \in \mathbb{S}} \mathbb{E}[1_{\{\alpha_s = j\}} 1_{\{\alpha_s^{\varepsilon h} = l\}} | \alpha_{k\eta} = i_1, \alpha_{k\eta}^{\varepsilon h} = i, X_{k\eta}, X_{k\eta}^{\varepsilon h}] \\
&= \sum_{l \in \mathbb{S}} q_{(i_1, i)(j, l)}(X_{k\eta}, X_{k\eta}^{\varepsilon h})(s - k\eta) + o(s - k\eta) \leq O(\eta). \tag{3.10}
\end{aligned}$$

Consequently, when  $k \neq 0$  the term  $I_{422k}$  in (3.8) can be estimated as follows

$$\begin{aligned}
&\mathbb{E} \sum_{i \in \mathbb{S}} \sum_{j \neq i} \int_{k\eta}^{(k+1)\eta} |b(X_{k\eta}^{\varepsilon h}, i) - b(X_{k\eta}^{\varepsilon h}, j)|^p 1_{\{\alpha_{k\eta}^{\varepsilon h} = i\}} 1_{\{\alpha_s = j\}} ds \\
&\leq C \mathbb{E} \sum_{i, i_1 \in \mathbb{S}} \sum_{j \neq i} \int_{k\eta}^{(k+1)\eta} [1 + |X_{k\eta}^{\varepsilon h}|^p] 1_{\{\alpha_{k\eta}^{\varepsilon h} = i, \alpha_{k\eta} = i_1\}} \\
&\quad \times \mathbb{E}[1_{\{\alpha_s = j\}} | \alpha_{k\eta} = i_1, \alpha_{k\eta}^{\varepsilon h} = i, X_{k\eta}, X_{k\eta}^{\varepsilon h}] ds \leq O(\eta^2). \tag{3.11}
\end{aligned}$$



Notice that when  $k = 0$ ,  $X_0^{\varepsilon h} = x$  and  $\alpha_0^{\varepsilon h} = \alpha$ . Thus we have

$$\begin{aligned} & \mathbb{E} \int_0^\eta |b(X_0^{\varepsilon h}, \alpha_0^{\varepsilon h}) - b(X_0^{\varepsilon h}, \alpha_s)|^p ds \\ &= \mathbb{E} \int_0^\eta \sum_{i \neq \alpha} |b(x, \alpha) - b(x, i)|^p 1_{\{\alpha_s = i\}} ds \\ &\leq \int_0^\eta \sum_{i \neq \alpha} |b(x, \alpha) - b(x, i)|^p [q_{\alpha i}(x)s + o(s)] ds \leq O(\eta^2). \end{aligned} \quad (3.12)$$

Combining the estimates (3.5)-(3.12), we obtain

$$I_4 \leq \sum_{k=0}^M C\eta^2 \leq C\eta. \quad (3.13)$$

In a similar way we have

$$I_5 \leq C\eta. \quad (3.14)$$

By (3.4), (3.13) and (3.14), we obtain

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\phi_t^\varepsilon|^p \right] \leq C \frac{\eta}{\varepsilon^p}.$$

For any  $\varepsilon > 0$  we choose  $\eta$  such that  $\eta < \varepsilon^q$ , where  $q > p$ . In this way we obtain the bound  $C\varepsilon^{q-p}$  which converges to zero as  $\varepsilon$  tends to zero. This completes the proof of the lemma. ■

**Lemma 3.5.** *Suppose that Hypothesis (H1) holds. Then for any  $h \in H$  and  $p \geq 2$ , we have*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^{\varepsilon h} - X_t|^p \right] \leq C\varepsilon^p.$$

*Proof* We write

$$\begin{aligned} X_t^{\varepsilon h} - X_t &= \int_0^t [b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_s, \alpha_s)] ds + \varepsilon \int_0^t \sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) h_s ds \\ &\quad + \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - \sigma(X_s, \alpha_s)] dW_s \\ &:= A(t) + B(t), \end{aligned}$$

where

$$\begin{aligned} A(t) &= \int_0^t [b(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - b(X_s^{\varepsilon h}, \alpha_s)] ds + \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - \sigma(X_s^{\varepsilon h}, \alpha_s)] dW_s \\ &\quad + \varepsilon \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s^{\varepsilon h}) - \sigma(X_s^{\varepsilon h}, \alpha_s)] h_s ds \end{aligned}$$

and

$$\begin{aligned} B(t) &= \int_0^t [b(X_s^{\varepsilon h}, \alpha_s) - b(X_s, \alpha_s)] ds + \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s) - \sigma(X_s, \alpha_s)] dW_s \\ &\quad + \varepsilon \int_0^t \sigma(X_s^{\varepsilon h}, \alpha_s) h_s ds. \end{aligned}$$

With this notation we can write

$$\begin{aligned} \sup_{0 \leq t \leq T} |X_t^{\varepsilon h} - X_t|^p &\leq C \left( \sup_{0 \leq t \leq T} |A(t)|^p + \sup_{0 \leq t \leq T} |B(t)|^p \right) \\ &= C\varepsilon^p \sup_{0 \leq t \leq T} |\phi_t^\varepsilon|^p + C \sup_{0 \leq t \leq T} |B(t)|^p. \end{aligned}$$

Applying Hölder's and Burkholder-Davis-Gundy's inequalities yields

$$\begin{aligned} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^{\varepsilon h} - X_t|^p \right] &\leq C\varepsilon^p \mathbb{E} \left[ \sup_{0 \leq t \leq T} |\phi_t^\varepsilon|^p \right] + C \int_0^T \mathbb{E} |X_s^{\varepsilon h} - X_s|^p ds \\ &\quad + C\varepsilon^p. \end{aligned}$$

By Lemma 3.4 and Gronwall's inequality, we obtain

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^{\varepsilon h} - X_t|^p \right] \leq C\varepsilon^p,$$

which completes the proof. ■

*Proof of Theorem 3.1* Let  $\psi_t^h$  be the solution of equation (3.1). It is easy to show that  $\mathbb{E} [\sup_{0 \leq t \leq T} |\psi_t^h|^p] \leq C$ , where  $C$  is a constant depending on  $T, x, h$  and  $p$ . We have

$$\begin{aligned} \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} - \psi_t^h &= \frac{1}{\varepsilon} \int_0^t [b(X_s^{\varepsilon h}, \alpha_s) - b(X_s, \alpha_s) - \varepsilon \nabla b(X_s, \alpha_s) \psi_s^h] ds \\ &\quad + \frac{1}{\varepsilon} \int_0^t \sum_{i=1}^d [\sigma_i(X_s^{\varepsilon h}, \alpha_s) - \sigma_i(X_s, \alpha_s) - \varepsilon \nabla \sigma_i(X_s, \alpha_s) \psi_s^h] dW_s^i \\ &\quad + \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s) - \sigma(X_s, \alpha_s)] h_s ds + \phi_t^\varepsilon. \end{aligned}$$

Using twice the mean valued theorem we have

$$\begin{aligned} &\frac{X_t^{\varepsilon h} - X_t}{\varepsilon} - \psi_t^h \\ &= \int_0^t \left[ \left( \int_0^1 \nabla b(X_s + \nu(X_s^{\varepsilon h} - X_s), \alpha_s) d\nu \right) \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \nabla b(X_s, \alpha_s) \psi_s^h \right] ds \\ &\quad + \int_0^t \sum_{i=1}^d \left[ \left( \int_0^1 \nabla \sigma_i(X_s + \nu(X_s^{\varepsilon h} - X_s), \alpha_s) d\nu \right) \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \nabla \sigma_i(X_s, \alpha_s) \psi_s^h \right] dW_s^i \\ &\quad + \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s) - \sigma(X_s, \alpha_s)] h_s ds + \phi_t^\varepsilon \\ &= \int_0^t \left( \int_0^1 \nabla b(X_s + \nu(X_s^{\varepsilon h} - X_s), \alpha_s) d\nu \right) \left( \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \psi_s^h \right) ds \\ &\quad + \int_0^t \sum_{i=1}^d \left( \int_0^1 \nabla \sigma_i(X_s + \nu(X_s^{\varepsilon h} - X_s), \alpha_s) d\nu \right) \left( \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \psi_s^h \right) dW_s^i + \phi_t^\varepsilon + \varphi_t^\varepsilon, \end{aligned}$$

where  $\phi_t^\varepsilon$  is defined before and

$$\begin{aligned}\varphi_t^\varepsilon &= \int_0^t [\sigma(X_s^{\varepsilon h}, \alpha_s) - \sigma(X_s, \alpha_s)] h_s ds \\ &\quad + \int_0^t \left( \int_0^1 \nabla b(X_s + \nu(X_s^{\varepsilon h} - X_s), \alpha_s) d\nu - \nabla b(X_s, \alpha_s) \right) \psi_s^h ds \\ &\quad + \int_0^t \sum_{i=1}^d \left( \int_0^1 \nabla \sigma_i(X_s + \nu(X_s^{\varepsilon h} - X_s), \alpha_s) d\nu - \nabla \sigma_i(X_s, \alpha_s) \right) \psi_s^h dW_s^i.\end{aligned}$$

By Hypothesis **(H<sub>2</sub>)** (this is the only place in the proof of the theorem we use this condition which is stronger than **(H<sub>1</sub>)**), we obtain

$$\begin{aligned}&\mathbb{E} \left[ \sup_{0 \leq s \leq t} \left| \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \psi_s^h \right|^p \right] \\ &\leq C \int_0^t \mathbb{E} \left| \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \psi_s^h \right|^p ds + C \mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s^{\varepsilon h} - X_s|^p \right] \\ &\quad + C \left( \mathbb{E} \sup_{0 \leq s \leq t} |X_s^{\varepsilon h} - X_s|^{2p} \right)^{1/2} \left( \mathbb{E} \sup_{0 \leq s \leq t} |\psi_s^h|^{2p} \right)^{1/2} + C \mathbb{E} \sup_{0 \leq s \leq t} |\phi_s^\varepsilon|^p.\end{aligned}$$

Using Gronwall's inequality and Lemmas 3.4 and 3.5, we obtain

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sup_{0 \leq s \leq t} \left| \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \psi_s^h \right|^p \right] = 0.$$

This implies that for any  $p \geq 2$ ,

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}_1 \left[ \sup_{0 \leq s \leq t} \left\| \frac{X_s^{\varepsilon h} - X_s}{\varepsilon} - \psi_s^h \right\|_{\mathbb{R}^n \otimes V}^p \right] = 0.$$

Now, let  $D_s X_t$  be the solution of the following equation:

$$D_s X_t = \sigma(X_s, \alpha_s) + \int_s^t \nabla b(X_r, \alpha_r) D_s X_r dr + \int_s^t \sum_{i=1}^d \nabla \sigma_i(X_r, \alpha_r) D_s X_r dW_r^i$$

for  $s \leq t$  and  $D_s X_t = 0$  for  $s > t$ . Then we can easily obtain that  $D^h X_t = \psi_t^h$  and  $DX_t \in L^{\infty-}(\Omega_1, H \otimes \mathbb{R}^n \otimes V)$ . The proof is complete. ■

Theorem 3.1 says that  $D^h X_t$  is a directional derivative of the solution  $X_t$  when we shift the Brownian motion with the deterministic function  $h$ . It is surprising to remark that the switching variable  $\alpha_t$  is not modified under this perturbation. As a consequence, we can show the following version of the chain rule.

**Theorem 3.6. (Chain rule)** *Assume that the condition **(H<sub>2</sub>)** holds. Then for any  $h \in H$ ,  $t \in [0, T]$  and  $p \geq 2$ , if  $f \in C_b^2(\mathbb{R}^n \times \mathbb{S})$ , we have*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) - f(X_t, \alpha_t)}{\varepsilon} - \nabla f(X_t, \alpha_t) D^h X_t \right|^p = 0.$$

Moreover,  $f(X_t, \alpha_t) \in \mathbb{D}^{1,\infty}(V)$  and  $Df(X_t, \alpha_t) = \nabla f(X_t, \alpha_t) DX_t$ .

*Proof* We can write

$$\begin{aligned} & \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) - f(X_t, \alpha_t)}{\varepsilon} - \nabla f(X_t, \alpha_t) D^h X_t \right|^p \\ & \leq C \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) - f(X_t, \alpha_t)}{\varepsilon} - \nabla f(X_t, \alpha_t) \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} \right|^p \\ & \quad + C \mathbb{E} \left| \nabla f(X_t, \alpha_t) \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} - \nabla f(X_t, \alpha_t) D^h X_t \right|^p. \end{aligned} \quad (3.15)$$

For the second term in (3.15), by Theorem 3.1 we have

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \nabla f(X_t, \alpha_t) \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} - \nabla f(X_t, \alpha_t) D^h X_t \right|^p = 0. \quad (3.16)$$

Consider now the first term in (3.15). For this term we make the following decomposition

$$\begin{aligned} & \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) - f(X_t, \alpha_t)}{\varepsilon} - \nabla f(X_t, \alpha_t) \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} \right|^p \\ & \leq C \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) - f(X_t^{\varepsilon h}, \alpha_t)}{\varepsilon} \right|^p \\ & \quad + C \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t) - f(X_t, \alpha_t)}{\varepsilon} - \nabla f(X_t, \alpha_t) \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} \right|^p. \end{aligned} \quad (3.17)$$

Similarly to the proof of Lemma 3.4, we can prove

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t^{\varepsilon h}) - f(X_t^{\varepsilon h}, \alpha_t)}{\varepsilon} \right|^p = 0. \quad (3.18)$$

Since  $f \in C_b^2(\mathbb{R}^n \times \mathbb{S})$  we have

$$\begin{aligned} & \mathbb{E} \left| \frac{f(X_t^{\varepsilon h}, \alpha_t) - f(X_t, \alpha_t)}{\varepsilon} - \nabla f(X_t, \alpha_t) \frac{X_t^{\varepsilon h} - X_t}{\varepsilon} \right|^p \\ & \leq C \mathbb{E} \frac{|X_t^{\varepsilon h} - X_t|^{2p}}{\varepsilon^p} \leq C \varepsilon^p. \end{aligned} \quad (3.19)$$

From (3.15)-(3.19) the theorem follows. ■

#### 4. SMOOTHNESS OF THE DENSITY

In this section we show that under suitable non degeneracy assumptions on the coefficients, for any  $t \in (0, T]$  the random vector  $X_t$  has a smooth density. To this end we first study the stochastic flow associated with equation (2.1) and then we show that the determinant of the Malliavin covariance matrix of  $X_t$  has finite negative moments of all orders for any  $t \in (0, T]$ , under a uniform Hörmander's condition.

**Definition 4.1.** Suppose that  $F(x, \alpha) : \Omega \rightarrow \mathbb{R}^n$  is a measurable function for all  $x \in \mathbb{R}^n$  and  $\alpha \in \mathbb{S}$ . We say that its gradient with respect to  $x$  exists (in mean square sense) if there is  $A(x, \alpha) : \Omega \rightarrow \mathbb{R}^{n^2}$  such that for any  $\xi \in \mathbb{R}^n$  we have

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \frac{F(x + \varepsilon \xi, \alpha) - F(x, \alpha)}{\varepsilon} - A(x, \alpha) \xi \right|^2 = 0.$$

We denote the gradient matrix  $A(x, \alpha)$  by  $\nabla F(x, \alpha)$ .

By Theorem 4.2 in [15] or by arguments similar to those used in the proof of Theorem 3.1, we can obtain the following results.

**Theorem 4.2.** *Assume the hypothesis  $(\mathbf{H}_2)$ . Let  $\{(X_{s,t}(x, \alpha), \alpha_{s,t}(x, \alpha)), t \geq s\}$  be the solution of equations (2.1) and (2.2), which starts from  $(x, \alpha)$  at time  $s$ . Then the gradient of  $X_{s,t}(x, \alpha)$  with respect to  $x$  (in mean square) exists. If we denote*

$$J_{s,t} := \nabla X_{s,t}(x, \alpha),$$

then

$$\begin{cases} dJ_{s,t} = \nabla b(X_t, \alpha_t) J_{s,t} dt + \sum_{i=1}^d \nabla \sigma_i(X_t, \alpha_t) J_{s,t} dW_t^i, & t \geq s \\ J_{s,s} = I, \end{cases} \quad (4.1)$$

where  $I$  is the  $n$ -dimensional identity matrix. Moreover,  $J_{s,t}$  is invertible and its inverse  $J_{s,t}^{-1}$  satisfies

$$\begin{cases} dJ_{s,t}^{-1} = -J_{s,t}^{-1} \left( \nabla b(X_t, \alpha_t) - \sum_{i=1}^d \nabla \sigma_i(X_t, \alpha_t) \nabla \sigma_i(X_t, \alpha_t) \right) dt \\ \quad - \sum_{i=1}^d J_{s,t}^{-1} \nabla \sigma_i(X_t, \alpha_t) dW_t^i, \\ J_{s,s}^{-1} = I. \end{cases} \quad (4.2)$$

The following lemma provides estimates on the  $L^p$  norm of the gradient of the solution and its inverse.

**Lemma 4.3.** *Assume that Hypothesis  $(\mathbf{H}_2)$  holds. Then for any  $p \geq 2$ , there exists a positive constant  $C$  depending only on  $T$  and  $p$  such that*

$$\mathbb{E} \left[ \sup_{s \leq t \leq T} (\|J_{s,t}\|^p + \|J_{s,t}^{-1}\|^p) \right] \leq C.$$

*Proof* From

$$J_{s,t} = I + \int_s^t \nabla b(X_r, \alpha_r) J_{s,r} dr + \int_s^t \nabla \sigma(X_r, \alpha_r) J_{s,r} dW_r$$

we obtain

$$\|J_{s,t}\|^p \leq C \left[ 1 + \left\| \int_s^t \nabla b(X_r, \alpha_r) J_{s,r} dr \right\|^p + \left\| \int_s^t \nabla \sigma(X_r, \alpha_r) J_{s,r} dW_r \right\|^p \right]$$

for any  $p \geq 2$ . Then by Hölder's inequality, we can write

$$\mathbb{E} \left[ \sup_{s \leq t \leq T} \left\| \int_s^t \nabla b(X_r, \alpha_r) J_{s,r} dr \right\|^p \right] \leq C \int_s^T \mathbb{E} \|J_{s,r}\|^p dr,$$

and by Burkholder-Davis-Gundy's inequality we have

$$\mathbb{E} \left[ \sup_{s \leq t \leq T} \left\| \int_s^t \nabla \sigma(X_r, \alpha_r) J_{s,r} dW_r \right\|^p \right] \leq C \int_s^T \mathbb{E} \|J_{s,r}\|^p ds.$$

Hence, we have

$$\mathbb{E} \left[ \sup_{s \leq t \leq T} \|J_{s,t}\|^p \right] \leq C + C \int_s^T \mathbb{E} \|J_{s,r}\|^p ds.$$

Then, Gronwall's inequality yields

$$\mathbb{E} \left[ \sup_{s \leq t \leq T} \|J_{s,t}\|^p \right] \leq C.$$

Similarly,

$$\mathbb{E} \left[ \sup_{s \leq t \leq T} \|J_{s,t}^{-1}\|^p \right] \leq C.$$

The proof of the lemma is now complete.  $\blacksquare$

Using the gradient of the flow  $J_{s,t}$  we can represent the Malliavin derivative  $DX_t$  as follows:

$$D_s X_t = J_{s,t} \sigma(X_s, \alpha_s), \quad 0 \leq s \leq t \leq T; \quad D_s X_t = 0, s > t.$$

Next, we shall study the Malliavin differentiability of  $J_{s,t}$ .

**Lemma 4.4.** *Suppose that Hypothesis  $(\mathbf{H}_2)$  holds. Then the following two statements hold:*

- (i) *For all  $0 \leq s \leq t \leq T$ ,  $J_{s,t} \in \mathbb{D}^{1,\infty}(\mathbb{R}^n \otimes \mathbb{R}^n \otimes V)$  and for any  $p \geq 2$ , there exists a positive constant  $C$  depending on  $T$ ,  $p$  and  $x$ , such that for all  $i = 1, \dots, d$  and  $r \in [0, T]$*

$$\mathbb{E} \left[ \sup_{s \leq t \leq T} \|D_r^i J_{s,t}\|^p \right] \leq C$$

- (ii) *For any  $t \leq T$ ,  $X_t \in \mathbb{D}^{2,\infty}(\mathbb{R}^n \otimes V)$  and for any  $p \geq 2$ , there exists a positive constant  $C$  depending on  $T$ ,  $p$  and  $x$ , such that for all  $i, j = 1, \dots, d$  and  $r, s \leq t$*

$$\mathbb{E} \|D_r^i (D_s^j X_t)\|^p \leq C.$$

*Proof* First, we prove (i). By equation (4.1), the definition of Malliavin derivative and the arguments used in the proof of Theorem 3.1, we obtain

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}_1 \left\| \frac{J_{s,t}(\omega_1 + \varepsilon \int_0^t h_s ds, \omega_2) - J_{s,t}(\omega_1, \omega_2)}{\varepsilon} - \langle DJ_{s,t}, h \rangle_H \right\|_{\mathbb{R}^n \otimes \mathbb{R}^n \otimes V}^p = 0$$

holds for any  $h \in H$  and any  $p \geq 2$ , where  $DJ_{s,t}$  satisfies the following equation for all  $r \leq t$

$$\begin{aligned} D_r^i J_{s,t} &= \nabla \sigma_i(X_r, \alpha_r) J_{s,r} 1_{\{s \leq r \leq t\}} + \int_{r \vee s}^t \left[ \nabla^2 b(X_u, \alpha_u) (D_r^i X_u, J_{s,u}) \right. \\ &\quad \left. + \nabla b(X_u, \alpha_u) D_r^i J_{s,u} \right] du + \int_{r \vee s}^t \sum_{k=1}^d \left[ \nabla^2 \sigma_k(X_u, \alpha_u) (D_r^i X_u, J_{s,u}) \right. \\ &\quad \left. + \nabla \sigma_k(X_u, \alpha_u) D_r^i J_{s,u} \right] dW_u^k, \end{aligned} \quad (4.3)$$

and for  $r > t$ ,  $D_r^i J_{s,t} = 0$ . For  $s \leq r \leq t$ , from the above identity we deduce the following estimates

$$\begin{aligned} |D_r^i J_{s,t}|^p &\leq C |\nabla \sigma_i(X_r, \alpha_r) J_{s,r}|^p + C \left| \int_r^t \left[ \nabla^2 b(X_u, \alpha_u)(D_r^i X_u, J_{s,u}) \right. \right. \\ &\quad \left. \left. + \nabla b(X_u, \alpha_u) D_r^i J_{s,u} \right] du \right|^p + C \left| \int_r^t \left[ \nabla^2 \sigma(X_u, \alpha_u)(D_r^i X_u, J_{s,u}) \right. \right. \\ &\quad \left. \left. + \nabla \sigma(X_u, \alpha_u) D_r^i J_{s,u} \right] dW_u \right|^p. \end{aligned}$$

Applying Hölder's and Burkholder-Davis-Gundy's inequality, we have

$$\begin{aligned} \mathbb{E} \left[ \sup_{r \leq t \leq T} |D_r^i J_{s,t}|^p \right] &\leq C |J_{s,r}|^p + C \int_r^T \mathbb{E}(|D_r^i X_u|^p \cdot |J_{s,u}|^p) du \\ &\quad + C \int_r^T \mathbb{E} |D_r^i J_{s,u}|^p du \\ &\leq C |J_{s,r}|^p + C \int_r^T (\mathbb{E} |D_r^i X_u|^{2p})^{1/2} \cdot (\mathbb{E} |J_{s,u}|^{2p})^{1/2} du \\ &\quad + C \int_r^T \mathbb{E} |D_r^i J_{s,u}|^p du. \end{aligned}$$

Hence, Lemma 4.3 and Gronwall's inequality yield

$$\mathbb{E} \left[ \sup_{r \leq t \leq T} |D_r^i J_{s,t}|^p \right] \leq C e^{C(T-r)}. \quad (4.4)$$

The case  $r < s$  can be handled in a similar way and hence the statement (i) is proved.

(ii) Note that for any  $j = 1, \dots, d$ ,

$$D_s^j X_t = J_{s,t} \sigma_j(X_s, \alpha_s) \text{ when } s \leq t \text{ and } D_s^j X_t = 0 \text{ when } s > t.$$

An application of chain rule, which can be easily established, yields

$$D_r^i (D_s^j X_t) = (D_r^i J_{s,t}) \sigma_j(X_s, \alpha_s) + J_{s,t} \nabla \sigma_j(X_s, \alpha_s) D_r^j X_s.$$

Then, Hölder's inequality gives

$$\begin{aligned} \mathbb{E} |D_r^i (D_s^j X_t)|^p &\leq C \mathbb{E} |(D_r^i J_{s,t}) \sigma_j(X_s, \alpha_s)|^p + C \mathbb{E} |J_{s,t} \nabla \sigma_j(X_s, \alpha_s) D_r^j X_s|^p \\ &\leq C (\mathbb{E} \|D_r^i J_{s,t}\|^{2p})^{1/2} [\mathbb{E}(1 + |X_s|^{2p})]^{1/2} \\ &\quad + C (\mathbb{E} \|J_{s,t}\|^{2p})^{1/2} (\mathbb{E} |D_r^j X_s|^{2p})^{1/2} \leq C, \end{aligned}$$

which implies (ii). ■

**Remark 4.5.** Following the same procedure as above we can prove that if Hypothesis  $(H_\infty)$  holds, then  $J_{s,t} \in \mathbb{D}^\infty(\mathbb{R}^n \otimes \mathbb{R}^n \otimes V)$  and  $X_t \in \mathbb{D}^\infty(\mathbb{R}^n \otimes V)$ .

We denote by  $(DX_t)^*$  the transpose of the random matrix  $DX_t$ . From the relation between  $DX_t$  and  $J_{s,t}$ , we have  $(DX_t)^*(r) = \sigma(X_r, \alpha_r)^* J_{r,t}^*$ . Then, the

Malliavin matrix  $M_t$  of the random vector  $X_t$  is defined by:

$$\begin{aligned} M_t &= \langle DX_t, (DX_t)^* \rangle_H = \int_0^t J_{s,t} \sigma(X_s, \alpha_s) \sigma(X_s, \alpha_s)^* J_{s,t}^* ds \\ &= J_{0,t} \int_0^t J_{0,s}^{-1} \sigma(X_s, \alpha_s) \sigma(X_s, \alpha_s)^* (J_{0,s}^{-1})^* ds J_{0,t}^* \\ &= J_{0,t} C_t J_{0,t}^*, \end{aligned}$$

where

$$C_t = \int_0^t J_{0,s}^{-1} \sigma(X_s, \alpha_s) \sigma(X_s, \alpha_s)^* (J_{0,s}^{-1})^* ds,$$

is the so-called reduced Malliavin matrix of  $X_t$ .

Our aim is to show that, under a suitable non degeneracy conditions on the coefficients, the Malliavin matrix  $M_t$  is invertible  $\mathbb{P}$ -a.s. and the determinant of its inverse has negative moments of all orders. The difficulty in our current situation is that the vector fields  $b$  and  $\sigma_1, \dots, \sigma_d$  depend on the Markov switching process  $\alpha_t$ . To overcome this difficulty we follow the following procedure inspired by [5].

For  $t \in [0, T]$  we define  $N_t := N([0, t], m_0(m_0 - 1)K)$ , so  $\{N_t, t \in [0, T]\}$  is a Poisson process with parameter  $m_0(m_0 - 1)K$ . Conditioned on the number of jumps of the Poisson process up to time  $t$ , that is,  $N_t = k$ , there exists a random interval  $[T_1, T_2]$  with  $0 \leq T_1 < T_2 \leq t$ , such that  $T_2 - T_1 \geq \frac{t}{k+1}$ . This implies that  $\alpha_t = \alpha_{T_1}$  for all  $t \in [T_1, T_2]$  (because that the jump times of  $\alpha_t$  are a subset of the jump times of  $N_t$ ). On this random time interval, we will apply the classical techniques of Malliavin calculus.

To this end we need the following stopping time version of Norris lemma.

**Lemma 4.6.** *Let  $\tau_1 : \Omega \rightarrow [0, T]$  be a stopping time and let  $\xi_1, \xi_2$  be two  $\mathcal{F}_{\tau_1}$ -measurable random variables. Suppose that  $\beta(t), \gamma(t) = (\gamma_1(t), \dots, \gamma_d(t))$  and  $u(t) = (u_1(t), \dots, u_d(t))$  are  $\mathcal{F}_t$ -adapted processes. For any  $t \geq \tau_1$ , set*

$$\begin{aligned} a(t) &= \xi_1 + \int_{\tau_1}^t \beta(s) ds + \sum_{i=1}^d \int_{\tau_1}^t \gamma_i(s) dW_s^i \\ Y(t) &= \xi_2 + \int_{\tau_1}^t a(s) ds + \sum_{i=1}^d \int_{\tau_1}^t u_i(s) dW_s^i \end{aligned}$$

and assume that for some  $p \geq 2$

$$\mathbb{E} \left( \sup_{\tau_1 \leq t \leq T} (|\beta(t)| + |\gamma(t)| + |a(t)| + |u(t)|)^p \right) < \infty. \quad (4.5)$$

Consider another stopping time  $\tau_2 : \Omega \rightarrow [0, T]$  satisfying  $\tau_1 < \tau_2$  and  $\tau_2 - \tau_1 \geq c$  almost surely for some constant  $c > 0$ . Then, for any  $q > 8$  and  $r > 0$  such that  $18r < q - 8$ , there exists  $\varepsilon_0 = \delta_0 c^{\gamma_0}$ , where the positive constants  $\delta_0$  and  $\gamma_0$  depend on  $p, q, r$  and  $T$ , such that for all  $\varepsilon \in (0, \varepsilon_0)$

$$\mathbb{P} \left\{ \int_{\tau_1}^{\tau_2} Y_t^2 dt < \varepsilon^q, \int_{\tau_1}^{\tau_2} (|a(t)|^2 + |u(t)|^2) dt \geq \varepsilon \right\} \leq \varepsilon^{rp}.$$

The proof is similar to the proof of Lemma 2.3.2 in [9] and we omit the details. Just remark that the lower bound  $c$  on the length of the random interval  $[\tau_1, \tau_2]$  is crucial in the proof.



We are going to impose a uniform Hörmander's condition on the coefficients. To formulate this condition we need some notation. Consider the following sets of vector fields:

$$\begin{aligned}\Sigma_0 &= \{\sigma_1, \dots, \sigma_d\}, \\ \Sigma_n &= \{[\sigma_k, V], k = 0, \dots, d, V \in \Sigma_{n-1}\}, \quad n \geq 1, \\ \Sigma &= \cup_{n=0}^{\infty} \Sigma_n,\end{aligned}$$

where  $\sigma_0 = b - \frac{1}{2} \sum_{i=1}^d (\nabla \sigma_i) \sigma_i$  and  $[V, G] = (\nabla G)V - (\nabla V)G$  denotes the Lie bracket between two vector fields  $V$  and  $G$ .

The following uniform Hörmander's condition requires that the vector space spanned by  $\{V(x, \alpha), V \in \Sigma\}$  is  $\mathbb{R}^n$  for all  $(x, \alpha) \in \mathbb{R}^n \times \mathbb{S}$  in a uniform way, where  $V_j(x, \alpha)$  denotes the vector obtained by freezing the variables  $x$  and  $\alpha$  in the vector field  $V_j$ .

**(UHC)** (Uniform Hörmander's condition) Condition **(H<sub>∞</sub>)** holds and there exists an integer  $j_0 \geq 0$  and a constant  $c > 0$  such that

$$\sum_{j=0}^{j_0} \sum_{V \in \Sigma_j} (v^* V(x, \alpha))^2 \geq c, \quad (4.6)$$

for all  $x \in \mathbb{R}^n$ ,  $\alpha \in \mathbb{S}$  and  $v \in \mathbb{R}^n$  with  $|v| = 1$ .

**Theorem 4.7.** *Assume that the uniform Hörmander's condition **(UHC)** holds. Then for all  $0 < t \leq T$  the Malliavin matrix  $M_t$  of the random vector  $X_t$  is invertible  $\mathbb{P}$ -a.s. and  $\det(M_t^{-1}) \in L^p(\Omega)$  for all  $p \geq 2$ . As a consequence, for any  $t \in (0, T]$ , the law of  $X_t$  is absolutely continuous with respect to Lebesgues measure and the density is smooth.*

*Proof* We recall that  $M_t = J_{0,t} C_t J_{0,t}^*$ . By Lemma 4.3 it suffices to prove that  $\det(C_t^{-1}) \in L^p(\Omega)$  for all  $p \geq 2$ .

Recall that  $\{N_t = N([0, t], m_0(m_0 - 1)K), t \in [0, T]\}$  is a Poisson process with parameter  $\lambda := m_0(m_0 - 1)K$ . For a fixed  $t \in (0, T]$ , conditioned on  $N_t = k$ , there exists a random interval  $[T_1, T_2]$  such that  $T_2 - T_1 \geq \frac{t}{k+1}$  and  $\alpha_t = \alpha_{T_1}$  for all  $t \in [T_1, T_2]$ .

We introduce the following sets of vector fields:

$$\begin{aligned}\Sigma'_0 &= \Sigma_0; \\ \Sigma'_n &= \{[\sigma_k, V], k = 1, \dots, d, V \in \Sigma'_{n-1}\}; \\ &\quad [\sigma_0, V] + \frac{1}{2} \sum_{j=1}^d [\sigma_j, [\sigma_j, V]], V \in \Sigma'_{n-1}\}, \quad n \geq 1; \\ \Sigma' &= \cup_{n=0}^{\infty} \Sigma'_n.\end{aligned}$$

We denote by  $\Sigma_n(x, \alpha)$  (resp.  $\Sigma'_n(x, \alpha)$ ) the subset of  $\mathbb{R}^n$  obtained by freezing the variable  $x, \alpha$  in the vector fields of  $\Sigma_n$  (resp.  $\Sigma'_n$ ). Clearly, the vector spaces spanned by  $\Sigma(x, \alpha)$  or by  $\Sigma'(x, \alpha)$  coincide. By Hypothesis **(UHC)**, there exists an integer  $j_0 \geq 0$  and a  $c > 0$  such that

$$\inf_{x \in \mathbb{R}^n} \inf_{\alpha \in \mathbb{S}} \sum_{j=0}^{j_0} \sum_{V \in \Sigma'_j} (v^* V(x, \alpha))^2 \geq c, \quad (4.7)$$

for all  $|v| = 1$ .

For all  $j = 0, 1, \dots, j_0$ , denote  $m(j) = 2^{-4j}$  and define

$$E_j = \left\{ \sum_{V \in \Sigma'_j} \int_{T_1}^{T_2} (v^* J_{0,s}^{-1} V(X_s, \alpha_s))^2 ds \leq \varepsilon^{m(j)} \right\}.$$

Clearly  $\{v^* C_t v \leq \varepsilon\} \subset E_0$ . Consider the decomposition

$$E_0 \subseteq (E_0 \cap E_1^c) \cup (E_1 \cap E_2^c) \cup \dots \cup (E_{j_0-1} \cap E_{j_0}^c) \cup F,$$

where  $F = E_0 \cap E_1 \cap \dots \cap E_{j_0}$ . Then for any unit vector  $v$  we have

$$\begin{aligned} \mathbb{P}\{v^* C_t v \leq \varepsilon | N_t = k\} &\leq \mathbb{P}(E_0 | N_t = k) \\ &\leq \mathbb{P}(F | N_t = k) + \sum_{j=0}^{j_0-1} \mathbb{P}(E_j \cap E_{j+1}^c | N_t = k). \end{aligned}$$

We are going to estimate each term in the above sum. This will be done in two steps.

*Step 1:* We can write

$$\mathbb{P}(F | N_t = k) \leq \mathbb{P}(F \cap G | N_t = k) + \mathbb{P}(G^c | N_t = k), \quad (4.8)$$

where  $G := \{\sup_{T_1 \leq s \leq T_2} \|J_{0,s}\| \leq \frac{1}{\varepsilon^\beta}\}$ ,  $0 < 2\beta < m(j_0)$ . First we claim that when  $\varepsilon$  is sufficiently small, the intersection  $F \cap G \cap \{N_t = k\}$  is empty. In fact, taking into account the estimate (4.7), on  $N_t = k$ , we have

$$\begin{aligned} &\sum_{j=0}^{j_0} \sum_{V \in \Sigma'_j} \int_{T_1}^{T_2} (v^* J_{0,s}^{-1} V(X_s, \alpha_s))^2 ds \\ &= \sum_{j=0}^{j_0} \sum_{V \in \Sigma'_j} \int_{T_1}^{T_2} \left( \frac{v^* J_{0,s}^{-1} V(X_s, \alpha_s)}{|v^* J_{0,s}^{-1}|} \right)^2 |v^* J_{0,s}^{-1}|^2 ds \geq \frac{tc\varepsilon^{2\beta}}{k+1}, \end{aligned} \quad (4.9)$$

because  $|v^* J_{0,s}^{-1}| \geq \frac{1}{\|J_{0,s}\|} \geq \varepsilon^\beta$ , and  $T_2 - T_1 \geq \frac{t}{k+1}$ . On the other hand, the left-hand side of (4.9) is bounded by  $(j_0 + 1)\varepsilon^{m(j_0)}$  on the set  $F$ . Thus  $F \cap G \cap \{N_t = k\} = \emptyset$ , provided  $\varepsilon < \varepsilon_1$ , where  $\varepsilon_1 = \left[ \frac{tc}{(k+1)(j_0+1)} \right]^{\frac{1}{m(j_0)-2\beta}}$ .

Now we consider the second term in (4.8). Using Chebychev inequality we obtain

$$\mathbb{P} \left( \sup_{T_1 \leq s \leq T_2} \|J_{0,s}\| \geq \varepsilon^{-\beta} | N_t = k \right) \leq \varepsilon^{p\beta} \mathbb{E} \left( \sup_{T_1 \leq s \leq T_2} \|J_{0,s}\|^p | N_t = k \right).$$

Taking into account that the Poisson random measure  $N$  is independent of the Brownian motion  $W$ , we can estimate the above conditional expectation using Burkholder-Davis-Gundy's inequality as in Lemma 4.3, and we obtain the estimate

$$\mathbb{P} \left( \sup_{T_1 \leq s \leq T_2} \|J_{0,s}\| \geq \varepsilon^{-\beta} | N_t = k \right) \leq C\varepsilon^{p\beta}. \quad (4.10)$$

*Step 2:* We shall bound the other terms in (4.8). For any  $j = 0, \dots, j_0 - 1$  we have

$$\begin{aligned}
\mathbb{P}(E_j \cap E_{j+1}^c | N_t = k) &= \mathbb{P} \left\{ \sum_{V \in \Sigma'_j} \int_{T_1}^{T_2} (v^* J_{0,s}^{-1} V(X_s, \alpha_s))^2 ds \leq \varepsilon^{m(j)}, \right. \\
&\quad \left. \sum_{V \in \Sigma'_{j+1}} \int_{T_1}^{T_2} (v^* J_{0,s}^{-1} V(X_s, \alpha_s))^2 ds > \varepsilon^{m(j+1)} | N_t = k \right\} \\
&\leq \sum_{V \in \Sigma'_j} \mathbb{P} \left\{ \int_{T_1}^{T_2} (v^* J_{0,s}^{-1} V(X_s, \alpha_{T_1}))^2 ds \leq \varepsilon^{m(j)}, \right. \\
&\quad \sum_{i=1}^d \int_{T_1}^{T_2} (v^* J_{0,s}^{-1} [\sigma_i, V](X_s, \alpha_{T_1}))^2 ds + \int_{T_1}^{T_2} (v^* J_{0,s}^{-1} ([\sigma_0, V] \\
&\quad \left. + \frac{1}{2} \sum_{i=1}^d [\sigma_i, [\sigma_i, V]])(X_s, \alpha_{T_1}))^2 ds > \frac{\varepsilon^{m(j+1)}}{n(j)} | N_t = k \right\},
\end{aligned}$$

where  $n(j)$  denotes the cardinality of the set  $\Sigma'_j$ . Consider the continuous semimartingale  $\{v^* J_{0,t}^{-1} V(X_t, \alpha_{T_1}), T_1 \leq t < T_2\}$ . Itô's formula yields for any  $s \in [T_1, T_2)$

$$\begin{aligned}
v^* J_{0,t}^{-1} V(X_t, \alpha_{T_1}) &= v^* J_{0,T_1}^{-1} V(X_{T_1}, \alpha_{T_1}) + \int_{T_1}^t v^* J_{0,s}^{-1} \sum_{i=1}^d [\sigma_i, V](X_s, \alpha_{T_1}) dW_s^i \\
&\quad + \int_{T_1}^t v^* J_{0,s}^{-1} \left\{ [\sigma_0, V] + \frac{1}{2} \sum_{i=1}^d [\sigma_i, [\sigma_i, V]] \right\} (X_s, \alpha_{T_1}) ds.
\end{aligned}$$

Notice that  $8m(j+1) < m(j)$  and also notice condition (4.5) holds for any  $p$  and the fact that the Poisson random measure  $N$  is independent of  $W$ . An application of Lemma 4.6 to the semimartingale  $Y_t = v^* J_{0,t}^{-1} V(X_t, \alpha_{T_1})$  with the stopping times  $T_1$  and  $T_2$  which satisfy  $T_2 - T_1 \geq \frac{t}{k+1}$  on the set  $N_t = k$  yields

$$\mathbb{P}(E_j \cap E_{j+1}^c | N_t = k) \leq \varepsilon^p \quad (4.11)$$

for any  $p \geq 2$ , and for  $\varepsilon < \varepsilon_0$ , where  $\varepsilon_0 = \delta_0 \left(\frac{t}{k+1}\right)^{\gamma_0}$ . The exponents  $\delta_0$  and  $\gamma_0$  only depend on  $p$  and  $T$ . Therefore, from (4.10) and (4.11) we obtain

$$\mathbb{P}\{v^* C_t v \leq \varepsilon | N_t = k\} \leq \varepsilon^p,$$

for any  $p \geq 2$ , and for  $\varepsilon < \min(\varepsilon_0, \varepsilon_1)$ . Then, following the steps of Lemma 2.3.1 in [9], we can obtain that

$$\mathbb{P}\left\{ \inf_{|v|=1} v^* C_t v \leq \varepsilon | N_t = k \right\} \leq \varepsilon^p$$

for all  $0 < \varepsilon \leq C_1 \left(\frac{t}{k+1}\right)^{C_2}$  and for all  $p \geq 2$ , where  $C_1$  and  $C_2$  are positive constants depending on  $p$ ,  $T$  and  $n$ . Consequently,

$$\begin{aligned} \mathbb{E}|\det(C_t)|^{-p} &\leq \mathbb{E}\left(\inf_{|v|=1} v^* C_t v\right)^{-np} \\ &\leq \sum_{k=0}^{\infty} \mathbb{P}(N_t = k) \mathbb{E}\left(\inf_{|v|=1} v^* C_t v\right)^{-np} |N_t = k \\ &\leq \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{\lambda} \left[ C_1 \left(\frac{t}{k+1}\right)^{C_2} + \frac{1}{C_1} \left(\frac{k+1}{t}\right)^{C_2} \right] < \infty. \end{aligned}$$

The proof is now complete. ■

## 5. BISMUT FORMULA FOR STATE DEPENDENT SWITCHING DIFFUSION

In this section, we prove a version of Bismut formula for state-dependent switching diffusions. This formula will be used to obtain the strong Feller property for the state dependent switching diffusion  $X_t$  in Section 7.

**Theorem 5.1.** *Suppose the condition (UHC) holds. Then for any  $f \in C_b^2(\mathbb{R}^n \times \mathbb{S})$ , we have*

$$\nabla P_t f(x, \alpha) = \mathbb{E} \left[ f(X_t, \alpha_t) \int_0^t \sigma(X_s, \alpha_s)^* J_{s,t}^* M_t^{-1} J_{0,t} dW_s \right],$$

where  $M_t = \int_0^t J_{s,t} \sigma(X_s, \alpha_s) \sigma(X_s, \alpha_s)^* J_{s,t}^* ds$  and the stochastic integral is interpreted in the Skorohod sense, that is,  $\int_0^t \sigma(X_s, \alpha_s)^* J_{s,t}^* M_t^{-1} J_{0,t} dW_s$  is the divergence of the process  $\{\sigma(X_s, \alpha_s)^* J_{s,t}^* M_t^{-1} J_{0,t} I_{[0,t]}(s), s \geq 0\}$ .

*Proof* For any  $\xi \in \mathbb{R}^n$  with  $|\xi| = 1$  let  $h^\xi = (DX_t)^* M_t^{-1} J_{0,t} \xi$ . Then we get

$$\begin{aligned} \langle DX_t, h^\xi \rangle_H &= \langle DX_t, (DX_t)^* M_t^{-1} J_{0,t} \xi \rangle_H \\ &= \langle DX_t, (DX_t)^* \rangle_H M_t^{-1} J_{0,t} \xi = J_{0,t} \xi. \end{aligned}$$

We claim that  $h^\xi \in \mathbb{D}^{1,p}(H \otimes V)$  for any  $p \geq 2$ . In fact, we have

$$\begin{aligned} D_s^i h^\xi &= (D_s^i (DX_t)^*) M_t^{-1} J_{0,t} \xi + (DX_t)^* M_t^{-1} (D_s^i J_{0,t}) \xi + (DX_t)^* (D_s^i M_t^{-1}) J_{0,t} \xi \\ &= (D_s^i (DX_t)^*) M_t^{-1} J_{0,t} \xi + (DX_t)^* M_t^{-1} (D_s^i J_{0,t}) \xi \\ &\quad - (DX_t)^* M_t^{-1} \left[ \langle D_s^i (DX_t), (DX_t)^* \rangle_H \right. \end{aligned} \tag{5.1}$$

$$\left. + \langle DX_t, D_s^i (DX_t)^* \rangle_H \right] M_t^{-1} J_{0,t} \xi. \tag{5.2}$$

By Lemma 4.3, Lemma 4.4 and Theorem 4.7 we obtain

$$\mathbb{E}_1 \|h^\xi\|_{H \otimes V}^p + \mathbb{E}_1 \|Dh^\xi\|_{H \otimes H \otimes V}^p \leq \mathbb{E} \|h^\xi\|_H^p + \sum_{i=1}^d \mathbb{E} \int_0^t \|D_s^i h^\xi\|_H^p ds < \infty.$$

Notice that  $h_s^\xi = \sigma(X_s, \alpha_s)^* J_{s,t}^* M_t^{-1} J_{0,t} \xi$ . Then, the derivative of the  $P_t f(x, \alpha)$  can be computed as follows

$$\begin{aligned}
\langle \nabla P_t f(x, \alpha), \xi \rangle &= \mathbb{E}(\nabla^\xi[f(X_t, \alpha_t)]) = \mathbb{E}(\nabla f(X_t, \alpha_t) J_{0,t} \xi) \\
&= \mathbb{E}(\nabla f(X_t, \alpha_t) \langle DX_t, h^\xi \rangle_H) = \mathbb{E}(\langle Df(X_t, \alpha_t), h^\xi \rangle_H) \\
&= \mathbb{E}_1(\langle Df(X_t, \alpha_t), h^\xi \rangle_{H \otimes V}) = \mathbb{E}_1(\langle f(X_t, \alpha_t), \delta(h^\xi) \rangle_V) \\
&= \mathbb{E}(f(X_t, \alpha_t) \delta(h^\xi)) \\
&= \mathbb{E} \left[ f(X_t, \alpha_t) \int_0^t \sigma(X_s, \alpha_s)^* J_{s,t}^* M_t^{-1} J_{0,t} \xi dW_s \right],
\end{aligned}$$

where the second and forth equalities follow from the chain rule, the sixth equality follows from the integration by parts formula, the stochastic integral is interpreted as in the sense of Skorohod. ■

## 6. THE EXISTENCE OF AND INVARIANT MEASURE

In this section we will consider the solution to the system of equations (2.1) and (2.3), where the time runs over  $[0, \infty)$  and we investigate the problem of existence of an invariant measure of the Markov process  $\{(X_t, \alpha_t), t \geq 0\}$ . We need to assume the following condition.

**(H2)** There exists two positive constants  $\lambda_1, \lambda_2$  such that for all  $x \in \mathbb{R}^n, i \in \mathbb{S}$ ,

$$2\langle x, b(x, i) \rangle + \sum_{k=1}^d |\sigma_k(x, i)|^2 \leq -\lambda_1 |x|^2 + \lambda_2.$$

**Theorem 6.1.** *Suppose that the conditions **(H1)** and **(H2)** hold. Then the solution  $\{(X_t, \alpha_t), t \geq 0\}$  admits at least one invariant measure.*

*Proof* By Itô's formula, we get

$$\begin{aligned}
|X_t|^2 &= |x|^2 + \int_0^t \left[ 2\langle X_s, b(X_s, \alpha_s) \rangle + \sum_{k=1}^d |\sigma_k(X_s, \alpha_s)|^2 \right] ds \\
&\quad + \sum_{k=1}^d \int_0^t 2\langle X_s, \sigma_k(X_s, \alpha_s) dW_s^k \rangle.
\end{aligned}$$

Taking the expectation on both sides and using Hypothesis **(H2)**, we can write

$$\begin{aligned}
\mathbb{E}|X_t|^2 &= |x|^2 + \mathbb{E} \int_0^t \left[ 2\langle X_s, b(X_s, \alpha_s) \rangle + \sum_{k=1}^d |\sigma_k(X_s, \alpha_s)|^2 \right] ds \\
&\leq |x|^2 - \lambda_1 \int_0^t \mathbb{E}|X_s|^2 ds + \lambda_2 t.
\end{aligned}$$

An application of Gronwall inequality yields

$$\begin{aligned}
\mathbb{E}|X_t|^2 &\leq e^{-\lambda_1 t} |x|^2 + \int_0^t e^{-\lambda_1(t-s)} \lambda_2 ds = e^{-\lambda_1 t} |x|^2 + \frac{\lambda_2}{\lambda_1} (1 - e^{-\lambda_1 t}) \\
&\leq |x|^2 + \frac{\lambda_2}{\lambda_1}.
\end{aligned} \tag{6.1}$$

Let  $f \in \mathcal{B}_b(\mathbb{R}^n \times \mathbb{S})$  and set for any  $T > 0$

$$\mu_T(f) := \frac{1}{T} \int_0^T P_t f(x, \alpha) dt.$$

Denote the ball centered at 0 with radius  $R$  by  $B_R = \{x \in \mathbb{R}^n; |x| \leq R\}$ . By (6.1) and Chebyshev's inequality, we have

$$\begin{aligned} \mu_T((B_R \times \mathbb{S})^c) &= \mu_T(B_R^c \times \mathbb{S}) = \frac{1}{T} \int_0^T P_t((x, \alpha), B_R^c \times \mathbb{S}) dt \\ &\leq \frac{1}{TR^2} \int_0^T \mathbb{E}|X_t(x, \alpha)|^2 dt \leq \frac{1}{R^2}(|x|^2 + \frac{\lambda_2}{\lambda_1}). \end{aligned}$$

Hence, for any  $\varepsilon > 0$ ,  $\sup_{T>0} \mu_T((B_R \times \mathbb{S})^c) < \varepsilon$  for  $R$  large enough. Then set of measures  $\{\mu_T, T > 0\}$  is tight. Therefore, there exist at least one measure  $\mu$  and a sequence  $T_n \rightarrow \infty$  such that  $\mu_{T_n} \rightarrow \mu$  weakly. By the classical Krylov-Bogoliubov theorem (see e.g. [4]),  $\mu$  is an invariant measure. ■

**Remark 6.2.** The condition **(H2)** is mainly used to obtain that  $\sup_{t \geq 0} \mathbb{E}|X_t|^2 < \infty$ , which ensures the transition probability family  $P_t((x, \alpha), \cdot), t \geq 0$  is tight. In fact, it is enough to show that there exists a constant  $p > 0$  such that  $\sup_{t \geq 0} \mathbb{E}|X_t|^p < \infty$ . So, the condition **(H2)** is somewhat strong, it can be weakened in some special case (see for example Lemma 3.2 in [1]).

## 7. ERGODICITY

It is well-known that in order to show the ergodicity of the Markov process  $\{(X_t, \alpha_t), t \geq 0\}$ , it is sufficient to show the existence and uniqueness of an invariant measure (see Theorem 3.2.6 in [4]). In section 6, we have already provided conditions for the existence of invariant measure. In this section we discuss the uniqueness of the invariant measure. For this we shall use the following well-known fact: the strong Feller property and irreducibility for the transition semigroup imply there exists at most one invariant measure (see Theorem 4.2.1 in [4]). In the following two subsections, we shall discuss the strong Feller property (subsection 7.1) and irreducibility (subsection 7.2) for the process  $\{(X_t, \alpha_t), t \geq 0\}$ .

**7.1. Strong Feller property.** To prove the strong Feller property, it suffices to prove that for any  $t > 0$  and for any bounded Borel measurable function  $f$  on  $\mathbb{R}^n \times \mathbb{S}$ ,  $P_t f(x, \alpha)$  is bounded and continuous in  $(x, \alpha)$ . Since  $\mathbb{S}$  is a finite set, it is sufficient to prove that for any  $\alpha \in \mathbb{S}$ ,  $P_t f(x, \alpha)$  is bounded and continuous with respect to  $x$ . We shall prove that this is a straightforward consequence of the Bismut formula.

**Theorem 7.1.** *Suppose that condition **(UHC)** holds. Then for any  $f \in \mathcal{B}_b(\mathbb{R}^n \times \mathbb{S})$ ,  $t > 0$ ,  $\alpha \in \mathbb{S}$ ,  $x \in \mathbb{R}^n$ , we have*

$$\lim_{y \rightarrow x} |P_t f(y, \alpha) - P_t f(x, \alpha)| = 0. \quad (7.1)$$

*Proof* Fix  $x \in \mathbb{R}^n$ ,  $\alpha \in \mathbb{S}$ ,  $t > 0$ . First, we consider the case that  $f \in C_b^2(\mathbb{R}^n \times \mathbb{S})$ . Applying the Bismut formula, we have

$$\begin{aligned} \|\nabla P_t f(x, \alpha)\| &= \sup_{|\xi|=1} |\langle \nabla P_t f(x, \alpha), \xi \rangle| \\ &\leq \sup_{|\xi|=1} \mathbb{E} \left| f(X_t, \alpha_t) \int_0^t h_s^\xi dW_s \right| \\ &\leq \|f\|_\infty \sup_{|\xi|=1} \mathbb{E}_1 \left\| \int_0^t h_s^\xi dW_s \right\|_V. \end{aligned}$$

As we have seen, since the process  $h_s^\xi$  is not adapted, the integral is Skorohod integral and it can be estimated as follows:

$$\begin{aligned} \mathbb{E}_1 \left\| \int_0^t h_s^\xi dW_s \right\|_V &\leq \left( \mathbb{E}_1 \left\| \int_0^t h_s^\xi dW_s \right\|_V^2 \right)^{1/2} \\ &= \mathbb{E}_1 \int_0^t \|h_s^\xi\|_{\mathbb{R}^d \otimes V}^2 ds + \mathbb{E}_1 \int_0^t \int_0^t \langle D_r h_s^\xi, D_s h_r^\xi \rangle_{\mathbb{R}^d \otimes \mathbb{R}^d \otimes V} dr ds \\ &\leq \mathbb{E}_1 \int_0^t \|h_s^\xi\|_{\mathbb{R}^d \times V}^2 ds + \mathbb{E}_1 \int_0^t \int_0^t \|D_r h_s^\xi\|_{\mathbb{R}^d \otimes \mathbb{R}^d \otimes V}^2 dr ds \\ &= \mathbb{E} \|h^\xi\|_H^2 + \sum_{i=1}^d \mathbb{E} \int_0^t \|D_s^i h^\xi\|_H^2 ds. \end{aligned}$$

Thus, we have

$$\|\nabla P_t f(x, \alpha)\| \leq C_x \|f\|_\infty,$$

where the constant  $C_x$  depends on the initial condition  $x$ . In fact, we also have that for any  $y \in B_r(x) = \{y \in \mathbb{R}^n : |y - x| \leq r\}$ , the following inequality holds

$$\sup_{y \in B_r(x)} \|\nabla P_t f(y, \alpha)\| \leq C_{x,r} \|f\|_\infty.$$

This implies easily for any  $|y - x| \leq 1$ ,

$$|P_t f(y, \alpha) - P_t f(x, \alpha)| \leq C_{x,1} \|f\|_\infty |y - x|.$$

Hence, the theorem holds for any  $f \in \mathcal{B}_b(\mathbb{R}^n \times \mathbb{S})$  by a standard argument. ■

**7.2. Irreducibility.** The transition semigroup  $\{P_t, t \geq 0\}$  is said to be irreducible, if for an arbitrary non-empty open set  $A$  in  $\mathbb{R}^n \times \mathbb{S}$  and for any  $(t, x, \alpha) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{S}$ ,  $P_t((x, \alpha), A) = P_t 1_A(x, \alpha) > 0$ . Since  $\mathbb{S}$  is a finite set, in order to show the irreducibility of the transition semigroup  $\{P_t, t \geq 0\}$ , it is sufficient to show that  $P_t((x, \alpha), B_\varepsilon(y) \times \{i\}) > 0$  for all  $t > 0$ ,  $(x, \alpha), (y, i) \in \mathbb{R}^n \times \mathbb{S}$ , and  $\varepsilon > 0$ .

For any  $k \in \mathbb{S}$ , let  $X_t^{(k)}(x)$  be the solution of the following stochastic differential equation in  $\mathbb{R}^n$ :

$$\begin{cases} dX_t^{(k)} = b(X_t^{(k)}, k)dt + \sum_{i=1}^d \sigma_i(X_t^{(k)}, k) dW_t^i, \\ X_0^{(k)} = x \in \mathbb{R}^n. \end{cases} \quad (7.2)$$

The corresponding transition semigroup is denoted by  $P_t^{(k)} f(x) = \mathbb{E} f(X_t^{(k)}(x))$ .

We need to assume the following condition:

**(H3)** For any  $k \in \mathbb{S}$ , the transition semigroup  $P_t^{(k)}$  of the diffusion process  $X_t^{(k)}$  is irreducible.

**Theorem 7.2.** *Assume that the conditions **(H1)** and **(H3)** hold and assume that  $q_{ij}(x) > 0$  for any  $x \in \mathbb{R}^n$  and  $i \neq j$ . Then, the transition semigroup  $\{P_t, t \geq 0\}$  is irreducible.*

*Proof* Put  $Y_t(x, \alpha) = (X_t(x, \alpha), \alpha_t(x, \alpha))^*$ ,  $\tilde{b}(x, \alpha) = (b(x, \alpha), 0)^*$ ,  $\tilde{\sigma}(x, \alpha) = (\sigma(x, \alpha), 0)^*$  and  $\tilde{g}(x, \alpha, z) = (0, g(x, \alpha, z))^*$ . Then we can rewrite equations (2.1) and (2.3) as the following form:

$$\begin{cases} dY_t = \tilde{b}(Y_t)dt + \tilde{\sigma}(Y_t)dW_t + \int_{[0, m_0(m_0-1)K]} \tilde{g}(Y_{t-}, z)N(dt, dz), \\ Y_0 = (x, \alpha) \in \mathbb{R}^n \times \mathbb{S}. \end{cases}$$

Recall the notation:

$$g(x, i, z) = \sum_{j \in \mathbb{S} \setminus i} (j - i) 1_{z \in \Delta_{ij}(x)}, \quad \forall i \in \mathbb{S},$$

where  $\Delta_{ij}(x)$  are the consecutive (with respect to the lexicographic ordering on  $\mathbb{S} \times \mathbb{S}$ ) left-closed, right-open intervals of  $\mathbb{R}_+$ , each having length  $q_{ij}(x)$ . As in [6],  $P_t$  satisfies the following equation.

$$P_t((x, \alpha), A) = e^{-tm_0(m_0-1)K} P_t^0((x, \alpha), A) \quad (7.3)$$

$$\begin{aligned} & + \sum_{j=1}^{m_0} \int_0^t \int_{\mathbb{R}^n} \int_{[0, m_0(m_0-1)K]} e^{-sm_0(m_0-1)K} \\ & \times P_s^0((x, \alpha), dy \times \{j\}) P_{t-s}((y, j + g(y, j, z)), A) dz ds, \end{aligned} \quad (7.4)$$

for any  $A \in \mathcal{B}(\mathbb{R}^n \times \mathbb{S})$ , where  $P_t^0((x, \alpha), A)$  is the transition probability for the following degenerate diffusion equation:

$$\begin{cases} d\tilde{Y}_t = \tilde{b}(\tilde{Y}_t)dt + \tilde{\sigma}(\tilde{Y}_t)dW_t, \\ \tilde{Y}_0 = (x, \alpha) \in \mathbb{R}^n \times \mathbb{S}. \end{cases} \quad (7.5)$$

By the relation between equations (7.2) and (7.5), we can easily get

$$P_t^0((x, \alpha), B_\varepsilon(y) \times \{i\}) = \begin{cases} P_t^{(\alpha)}(x, B_\varepsilon(y)) & \text{if } i = \alpha \\ 0 & \text{if } i \neq \alpha. \end{cases} \quad (7.6)$$

So, by the condition **(H3)** and (7.4), we have

$$\begin{aligned} P_t((x, \alpha), B_\varepsilon(y) \times \{\alpha\}) & \geq e^{-tm_0(m_0-1)K} P_t^0((x, \alpha), B_\varepsilon(y) \times \{\alpha\}) \\ & = e^{-tm_0(m_0-1)K} P_t^{(\alpha)}(x, B_\varepsilon(y)) > 0, \end{aligned}$$

for any  $t > 0, (x, \alpha) \in \mathbb{R}^n \times \mathbb{S}, y \in \mathbb{R}^n, \varepsilon > 0$ . Now it remains to show that  $P_t((x, \alpha), B_\varepsilon(y) \times \{i\}) > 0$ , for  $i \neq \alpha$ . Applying (7.4) again yields

$$\begin{aligned} P_{t-s}((x, \alpha), A) & = e^{-(t-s)m_0(m_0-1)K} P_{t-s}^0((x, \alpha), A) \\ & + \sum_{j=1}^{m_0} \int_0^{t-s} \int_{\mathbb{R}^n} \int_{[0, m_0(m_0-1)K]} e^{-s_1 m_0(m_0-1)K} \\ & \times P_{s_1}^0((x, \alpha), dy \times \{j\}) P_{t-s-s_1}((y, j + g(y, \alpha, z)), A) dz ds_1. \end{aligned}$$



Combining this with (7.4) we obtain

$$\begin{aligned}
P_t((x, \alpha), B_\varepsilon(y) \times \{i\}) &\geq e^{-tm_0(m_0-1)K} \\
&\times \sum_{j=1}^{m_0} \int_0^t \int_{\mathbb{R}^n} \int_{[0, m_0(m_0-1)K]} P_s^0((x, \alpha), dy_1 \times \{j\}) \\
&\times P_{t-s}^0((y_1, j + g(y_1, j, z)), B_\varepsilon(y) \times \{i\}) dz ds \\
&\geq e^{-tm_0(m_0-1)K} \\
&\times \int_0^t \int_{\mathbb{R}^n} \int_{\Delta_{\alpha i}(y_1)} P_s^0((x, \alpha), dy_1 \times \{\alpha\}) \\
&\times P_{t-s}^0((y_1, \alpha + g(y_1, \alpha, z)), B_\varepsilon(y) \times \{i\}) dz ds \\
&= e^{-tm_0(m_0-1)K} \\
&\times \int_0^t \int_{\mathbb{R}^n} \int_{\Delta_{\alpha i}(y_1)} P_s^{(\alpha)}(x, dy_1) P_{t-s}^{(i)}(y_1, B_\varepsilon(y)) dz ds.
\end{aligned}$$

Then  $P_t((x, \alpha), B_\varepsilon(y) \times \{i\}) > 0$  by Hypothesis **(H3)** and  $\Delta_{\alpha i}(y_1) > 0$  for any  $y_1 \in \mathbb{R}^n$ ,  $\alpha \neq i$ . The proof is complete. ■

**Remark 7.3.** Condition **(H3)** and  $q_{ij}(x) > 0$  seem natural. In fact, it guarantees that the continuous component  $X_t$  is irreducible for a fixed value  $\alpha_t = k$ , and, on the other hand, the hypothesis  $q_{ij}(x) > 0$  ensures that the discrete component  $\alpha_t$  is a irreducible Markov chain. Meanwhile, the Stroock-Varadhan's support theorem can be used to verify the condition **(H3)** (see Corollary 6.2 in [11]).

## REFERENCES

- [1] G. Basak, A. Bisi, M. Ghosh, *Stability of a Random Diffusion with Linear Drift*, J. Math. Anal. Appl. 202 (1996) 604-622.
- [2] T. Cass, *Smooth densities for solutions to stochastic differential equations with jump*, Stochastic Process. Appl. 119 (2009) 1416-1435.
- [3] M. Chen, *From Markov Chains to Non-equilibrium Particle Systems*, World Scientific, Singapore, 2004.
- [4] G. Da Prato and J. Zabczyk, *Ergodicity for infinite dimensional systems*, Cambridge University Press, 1996.
- [5] B. Forster, E. Lütkebohmert and J. Teichmann, *Absolutely continuous laws of jump-diffusions in finite and infinite dimensions with applications to mathematical finance*, SIAM J. Math. Anal. 40(5) (2009) 2132-2153.
- [6] C. Li, Z. Dong, R. Situ, *Almost sure stability of linear stochastic differential equations with jumps*, Probab. theory Related Fields 123 (2002) 121-155.
- [7] P. Malliavin, *Stochastic analysis*, Springer-Verlag, Berlin, 1997.
- [8] J. Norris, *Simplified Malliavin calculus*, Seminaire de probabilités (Strasbourg), 20 (1986), 101-130.
- [9] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer, Berlin, 2006.
- [10] Di Nunno, G.; Øksendal, B. and Proske, F. Malliavin calculus for Lévy processes with applications to finance. Universitext. Springer, Berlin, 2009.
- [11] L. Rey-Bellet, *Ergodic properties of Markov processes*, Open quantum systems. II, 1-39, Lecture Notes in Math., 1881, Springer, Berlin, 2006.
- [12] F. Xi, *Feller property and exponential ergodicity of diffusion process with state-dependent switching*, Science in China Series A: Mathematics, 51(3) (2008), 329-342.
- [13] F. Xi and G. Yin, *Jump-diffusions with state-dependent switching: existence and uniqueness, Feller property, linearization, and uniform ergodicity*, Science in China Series A: Mathematics, 54(12) (2011), 2651C2667.
- [14] F. Xi and L. Zhao, *On the stability of diffusion processes with state-dependent switching*, Science in China Series A: Mathematics, 49(9) (2006), 1258-1274.

- [15] G. Yin and C. Zhu, *Properties of solutions of stochastic differential equations with continuous-state-dependent switching*, J. Differential Equations 249 (2012) 2409-2439.
- [16] C. Zhu and G. Yin, *On the strong Feller, recurrence, and weak stabilization of regime-switching diffusion*, SIAM J. Control Optim. 48 (2009) 2003-2031.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF KANSAS, LAWRENCE, KANSAS, 66045  
*E-mail address*: hu@math.ku.edu

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF KANSAS, LAWRENCE, KANSAS, 66045  
*E-mail address*: nualart@math.ku.edu

SCHOOL OF MATHEMATICAL SCIENCES, NANKAI UNIVERSITY, TIANJIN 300071, CHINA., AND,  
SCHOOL OF MATHEMATICS AND STATISTICS, JIANGSU NORMAL UNIVERSITY, XUZHOU 221116, CHINA  
*E-mail address*: xbsun@mail.nankai.edu.cn, ycxie@jsnu.edu.cn